

デジタルアーカイブの構築・共有・活用ガイドライン

平成29年4月

デジタルアーカイブの連携に関する
関係省庁等連絡会・実務者協議会
(事務局:内閣府知的財産戦略推進事務局)



このガイドラインは、[クリエイティブ・コモンズ 表示 4.0 国際ライセンス](https://creativecommons.org/licenses/by/4.0/)の下に提供されています。

目次

はじめに.....	1
1. 我が国として目指すべきデジタルアーカイブ推進の方向性.....	4
2. デジタルアーカイブの整備に当たって.....	7
(1) メタデータの整備.....	8
(2) サムネイル/プレビューの作成.....	9
(3) デジタルコンテンツの作成・収集.....	9
(4) 長期アクセスの保証のために.....	11
3. データを共有するに当たって.....	14
(1) 公開ポリシーの考え方.....	14
(2) 二次利用条件の表示方法.....	15
(3) 望ましい利用条件(オープン化の推進).....	17
(4) 利用条件表示の検討に当たっての留意点.....	19
(5) データ共有の方法.....	20
4. データを活用するに当たって.....	24
(1) データの活用における留意点.....	25
(2) 付加価値情報の付与.....	25
(3) 情報間の関連付け.....	26
(4) 活用の結果できた成果物の還元.....	27
(5) 活用のためのコミュニティ形成.....	28
おわりに.....	29
(補足資料).....	30
●用語集.....	30
●利用条件表示一覧.....	32
(参考資料).....	33
●確認すべき標準・ガイドライン等.....	33
●活用できる表形式のデータとは?.....	39
●よくある質問.....	40
●ガイドライン早見表.....	42

はじめに

(本ガイドラインの背景)

ビッグデータ、IoT、3Dデータ、人工知能(AI)など、情報技術の革新は目覚ましく、多種多様なデジタル情報資源は、新たな社会的・文化的・経済的価値を様々に顕在化させ、アナログの世界を中心としていたこれまでの業務・サービスを一変させている。デジタル情報資源の効率的、効果的な活用は、博物館・美術館・図書館・文書館等の文化的活動はもちろん、大学、研究機関、学協会等の学術・研究活動や、シチズンサイエンスなどの市民活動、産業界における新ビジネス事業、教育や観光、地方創生といった国・地方自治体の行政に至るまで、コンテンツの保有者にとって非常に重要な課題となっている。

また、国が進めるオープンデータ戦略¹や、オープンサイエンスの推進²等の取組をはじめ、最近成立した「官民データ活用推進基本法」の趣旨や方向性に照らしても、様々なデジタル情報資源が、二次利用を促進する形でオープンに提供され、広く流通することが望まれる。様々な立場の人が多様な目的で活用できるデジタル情報資源が増えることで、新たなイノベーションが生まれ、社会が活性化する。多種多様なデジタル情報資源を抱える「デジタルアーカイブ」においても、このような成果につながる形での活用が求められている。

(本ガイドラインの対象者と目的)

本ガイドラインは、博物館・美術館、図書館、文書館といった文化的施設に加えて、大学・研究機関、企業、市民団体、官公庁・地方公共団体などの有形・無形の様々なコンテンツを保有する機関・団体等を「アーカイブ機関」³(コンテンツの保有者であれば、誰でも該当する。)と呼ぶこととし、まずはアーカイブ機関を対象に、業務にもサービスにも役立つデジタル情報資源の整備・運用方法を示した。

また、本ガイドラインは、これらアーカイブ機関に加えて、一般ユーザ、IT技術者、クリエイターなど、デジタルアーカイブ上にある様々なデジタル情報資源を利活用する機関・団体、個人も「活王者」として対象にしている。デジタルアーカイブの利活用に当たって、活王者はどのような点に留意すべきで何に取り組むべきか、についても示した。

さらに、アーカイブ機関と活王者との間に存在し、分野・地域ごとのコミュニティにおける

¹ <http://www.kantei.go.jp/jp/singi/it2/index.html>

http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyou/opendata/

² <http://www8.cao.go.jp/cstp/sonota/openscience/index.html>

³ 本ガイドラインでは、広い概念での記録機関という意味に捉え直し、コンテンツを保有している機関全般を対象とする。具体的には、補足資料「用語集」を参照のこと。

デジタル情報資源の共有化を促す「つなぎ役」⁴についても、本ガイドラインの中で取り組むべき事項を示している。

それぞれの機関・団体・個人が本ガイドラインに沿うことによって、我が国のデジタル情報資源が一層豊かになると同時に、アーカイブ機関自らもその恩恵を最大限に受けられるようになることを目指す。

なお、本ガイドラインの目的を深く理解するには、同時にとりまとめた「我が国におけるデジタルアーカイブ推進の方向性」⁵も併せて参照されたい。

(本ガイドラインにおける用語の整理)

本ガイドラインにおける「デジタルアーカイブ」は、様々なデジタル情報資源を収集・保存・提供する仕組みの総体を指す。そこで扱うデジタル情報資源は、資料・作品のデジタル化データやボーンデジタルの写真・動画等の「デジタルコンテンツ」はもちろん、アナログ媒体の資料・作品等も「コンテンツ」に含まれるものとした上で、「コンテンツ」に係る情報を記述した「メタデータ」、コンテンツの縮小版や部分表示である「サムネイル/プレビュー」など、関連するデジタルデータ全てを対象とする。また、本ガイドラインにおけるこれらの用語の関係は、次に示す図1のように三層構造に分けて扱う。なお、「メタデータ」という用語は、そもそもデータに関するデータのことであり、ここでいうメタデータとサムネイル/プレビュー、更には用語を統制するための語彙等も含むものとして広く定義されることもあるが、本ガイドラインは、図1に示すように、コンテンツの内容や所在等の情報を記述したデータを「メタデータ」と呼ぶこととする。

アーカイブ機関やつなぎ役なども含め、本ガイドラインで使われる主要な用語の意味するところに関しては、補足資料の用語集を参照されたい。

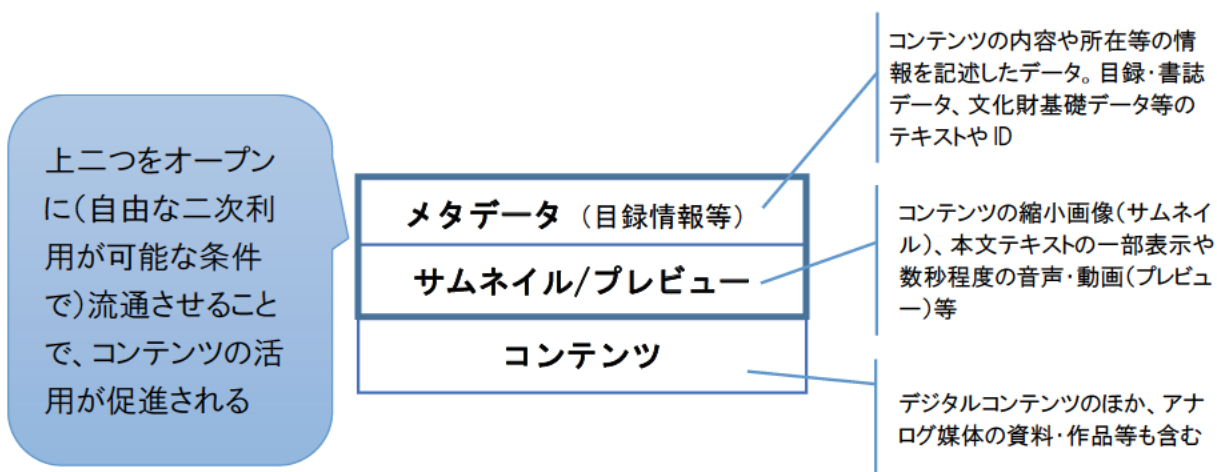


図1 デジタルアーカイブ連携における流通単位

⁴ Europeana でいうところのアグリゲーター、DPLA でいうところのハブに相当する役割・機能を果たす機関を想定する。具体的には、補足資料「用語集」を参照のこと。

⁵ http://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/houkokusho.pdf

(本ガイドラインの構成)

1章で我が国が目指すべきデジタルアーカイブの方向性を共有したうえで、2章でアーカイブ機関がデジタルアーカイブの整備を進めるに当たって行うべき事項を、3章ではそれらの整備されたデータを分野・地域のコミュニティ内やその他の分野・領域間において共有する場合に、アーカイブ機関やつなぎ役が行うべき事項を示している。4章では、アーカイブ機関に限らず行政や個人も含めた様々な主体がデジタルアーカイブのデータを活用すると想定し、その活用において求められる事項を示している。

各章ごとに、想定している対象者と重要度は次の通りである。

なお、4章はデジタルアーカイブの意義を具体的にイメージできる内容でもあるため、最初に読むことで理解が深まる可能性がある。

表1 目次ごとの対象者の参照箇所

【凡例】 ◎:必須情報 ○:参照情報

		アーカイブ機関	つなぎ役	活用者
1章	我が国として目指すべきデジタルアーカイブ推進の方向性	◎	◎	◎
2章	(1) メタデータの整備	◎	○	
	(2) サムネイル/プレビューの作成	◎	○	
	(3) デジタルコンテンツの作成・収集	◎	○	
	(4) 長期アクセスの保証のために	◎	◎	
3章	(1) 公開ポリシーの考え方	◎	○	
	(2) 二次利用条件の表示方法	◎	◎	
	(3) 望ましい利用条件(オープン化の推進)	◎	◎	
	(4) 利用条件表示の検討に当たっての留意点	◎	◎	
	(5) データ共有の方法	○	◎	
4章	(1) データの活用における留意点		○	◎
	(2) 付加価値情報の付与		○	○
	(3) 情報間の関連付け		◎	○
	(4) 活用の結果できた成果物の還元		○	◎
	(5) 活用のためのコミュニティ形成	○	◎	○

また、本ガイドライン末尾には、補足資料として「用語集」と「利用条件表示一覧」を付した。参考資料としては、本ガイドラインの理解を助ける技術的な情報を含む「確認すべき標準・ガイドライン等」「活用できる表形式のデータとは？」のほか、「よくある質問」「ガイドライン早見表」といった本ガイドラインの内容を簡単に確認できる資料を付しているの
でこちらも参考にされたい。

1. 我が国として目指すべきデジタルアーカイブ推進の方向性

(デジタルアーカイブ社会)

いま、様々な分野・地域における多様なシーンにおいて、デジタル情報資源が社会・文化・学術の基盤を支えている。できる限り効率的・効果的にこれらの資源を社会全体で有効に活用するためには、我が国では以下に示す各領域での取組が求められる(次ページの図2参照)。

〔保存・共有領域〕

「アーカイブ機関」は、自ら保有する、アナログも含めた各種コンテンツのメタデータの整備を行う。また、アナログ媒体の資料・作品のデジタル化や、ポーンデジタルの写真や動画等の収集によってデジタルコンテンツを拡充する。その際、サムネイル/プレビューの作成にも留意する。こうした活動は、誰がどのようなコンテンツを保有しているかを把握でき、資源の共有に役立つだけではない。活用できるデジタル情報資源が増えることは、自らの業務・サービスの向上にもつながる。

「つなぎ役」は、分野・地域のコミュニティをまとめ、各アーカイブ機関が提供する資源の共有化を図る役割を果たす。分野内・外におけるメタデータの共有化に加え、メタデータの標準化が重要な取組となる。標準化には、情報をつなぐ鍵となる要素、例えば、固有名情報、地理情報や年代・時間情報といった用語を統制することも含まれる。(標準化は、活用領域でも重要であり、本ガイドラインでは4章の活用箇所に記載している。)また、各アーカイブ機関が取り組むメタデータ整備やデジタルコンテンツ作成への支援、長期アクセス保証のための基盤提供などの役割を担うことも想定される。

〔活用領域〕

「国の分野横断統合ポータル」は、我が国が保有する様々なコンテンツへの効果的なナビゲーションを提供する。つまり、保存・共有領域にある多様な資源とその「活用者」とのつなぎ役を果たす。

デジタルアーカイブの「活用者」は、保存・共有領域でオープンになったデジタル情報資源に関して、その価値を一層高める方法で利用や活用を進める。例えば、特定のテーマのポータルやアプリを提供したり、複数の分野・地域をまたいだ情報間のつながりを生み出し、データに付加価値をつけるといった取組である。さらに活用者は、自らの成果をデータ提供者(アーカイブ機関やつなぎ役)に還元していくことによって、保存・共有領域の充実にも貢献する。

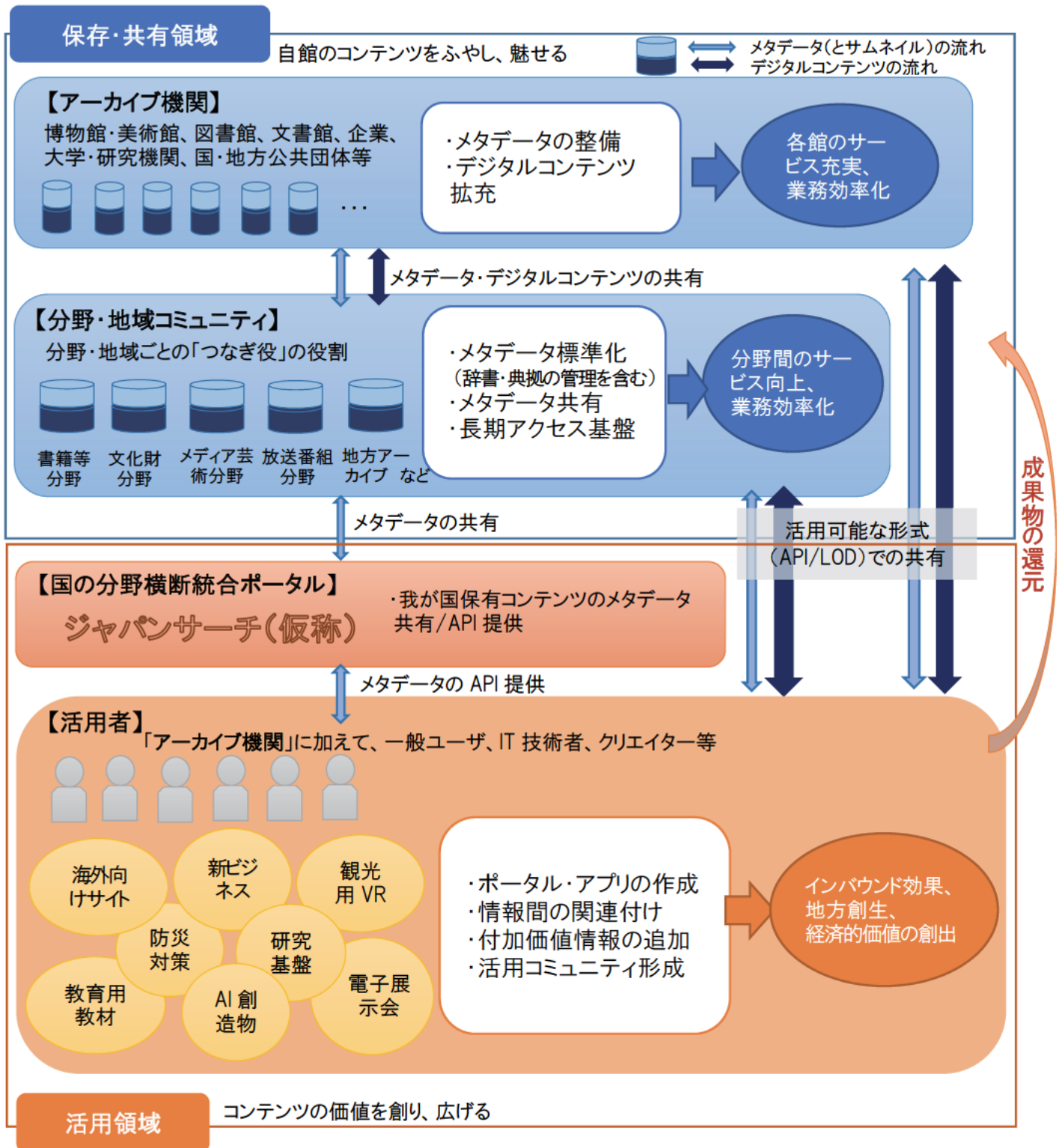


図 2 デジタルアーカイブの共有と活用のために

こうした領域ごとの取組によって、我が国のデジタル情報資源が効率的に生み出され、国全体として有効に活用されていくことを目指すものとする。詳細は前述の「我が国のデジタルアーカイブの推進の方向性」を参照のこと。

この全体の流れにおいて、各アーカイブ機関は、それぞれの立場や取組においてデジタルアーカイブに関わることになる。アーカイブ機関の多くは、コンテンツを保有してそれらを整備し提供すると同時に、他機関が提供する資源の活用者にもなりうる。例えば、自ら保有するデジタルコンテンツに加えて、他機関のデータも活用した展示会の開催などが考えられる。

また、ある程度の規模のアーカイブ機関は、同じ分野又は同じ圏域の機関のつなぎ役を兼ねることもありうる。例えば、都道府県立図書館は、市区町村立図書館のつなぎ役にもなりえるし、地域アーカイブを主導するなどしてその地域の博物館・美術館・文書館・大学・企業等のつなぎ役にもなりうる。アーカイブ機関ではないが、つなぎ役の役割の一部を国が担う場合もありうる。文化財分野では、文化庁の「文化遺産オンライン」がつなぎ役の一例である。

アーカイブ機関やデジタルアーカイブの活用者は、自らの関わりが、デジタルアーカイブ社会におけるどの領域のどの段階になるかを確認しつつ、以下の2～4章に示す取組を行っていくことが求められる。

2. デジタルアーカイブの整備に当たって

アーカイブ機関が、(利用提供していないものも含め)保有する有形・無形のコンテンツのメタデータやサムネイル/プレビューを整備して自らのコンテンツの管理を行うことは、日々の業務運営はもちろん、災害時の被害状況の把握などにも資するため、重要かつ必要なことである。また、有形のアナログ資料・収蔵品をデジタル化したり、ポーンデジタルのデータを収集したりして、デジタルコンテンツを拡充することは、情報技術を利用した効果的なサービスを展開するのに、欠かせないことである。ホームページでの発信や展示会等での利用など、自らのデジタル情報資源の最大の活用者は、結局のところ、その機関自身である。活用可能なデジタルコンテンツが増えることはアーカイブ機関自身にとって非常に重要であると同時に、デジタルアーカイブの推進にもつながることで、我が国全体にとっても大きな効果を持つ。

アーカイブ機関は、より効率的に業務運営を行い、魅力あるサービスを提供できるよう、本章に示す点に留意してデジタル情報資源を整備することが求められる。



図3 デジタルアーカイブの自館でのメリット(例)

(1) メタデータの整備

(メタデータの整備方法)

自らが保有するコンテンツのメタデータを整備して共有する方法には、自らが用意したシステムを用いるほか、分野や地域のコミュニティが提供する既存のプラットフォームを利用する方法がある。また、これらのシステム基盤を利用せずに、参考資料「活用できる表形式のデータとは？」に示す要件を満たす形の CSV 等のデータを整備する方法もある。誰がどんな内容のコンテンツを持っているかを示すメタデータが整備され共有されることは、同じコミュニティの他機関やデータの活用者にとって有益な情報提供となる。

メタデータの整備においては、次の 5 つの項目は共有や再利用のために特に重要であり、判明している場合には、必須の情報として記述することが求められる。

- ・タイトル(ラベル)
- ・作者(人物)
- ・日付(時代)
- ・場所
- ・管理番号(表内で重複しない恒久的な識別子)

作者や日付、場所についての項目は、不明の場合は不明でもよい。また、分野によって、これらが意味する情報が異なる場合がある。例えば、書籍であれば著者、出版年、出版地になるが、美術作品であれば作者、作成時代、所蔵場所といった情報が考えられる。このように分野によって対象や記述の仕方が異なるものもあるが、特に共有・公開を目的とするメタデータは、基本的には複雑にせずシンプルで一貫した記述がよい。各分野における事情を考慮した主要な標準については、参考資料「確認すべき標準・ガイドライン等」を参照し、管理や長期保存の必要に応じて用いることがよい。管理番号となる識別子の詳細は、「2(4)長期アクセスの保証のために(識別子の付与方法)」を参照のこと。また、デジタルコンテンツ作成・収集時のメタデータの整備に関する留意点は「2(3)デジタルコンテンツの作成・収集(デジタルコンテンツの蓄積)」を参照のこと。

このほか、デジタルアーカイブのデータ共有のためには、コンテンツの権利情報や二次利用条件といった情報もメタデータの整備において求められる。これらについては「3(2)二次利用条件の表示方法」を参照のこと。また、地理情報、時間情報などデータの活用者にとって有益なメタデータ項目については、「4(3)情報間の関連付け」を参照のこと。

(メタデータの整備に当たっての留意点)

アーカイブ機関は、可能な限り、時間の経過によって変化しないものをメタデータの必須項目として選びつつ、継続的にメンテナンスしていくことが求められる。

アーカイブ機関は、更新履歴、変更の差分、データ作成者の明示など、メタデータの信頼性を高める情報提供にも努めることがよい。

このほか、国際的な共有を考えた場合に、海外の活用者や日本国内の外国人のための多言語化(特に英語・ローマ字表記)など、メタデータの整備段階で留意することが望まれる事項がある。それについては「3(5)データ共有の方法」を参照のこと。

(2) サムネイル/プレビューの作成

対象を判別・同定する際の正確性を高めるため、メタデータの内容を充実させ、記述する項目を増やすという方法もあるが、主な特徴を示すコンテンツの要約物や一部分の表示によってメタデータの情報を補う方法もある。資料・作品中の本文テキストの一部（例えば、冒頭の数行だけ）をテキストデータとして入力することでもよいし、コンテンツの縮小画像（サムネイル）や音声・動画の部分抽出（プレビュー）といったものを作成することでもよい。

サムネイルやプレビューは、次項の「2(3)デジタルコンテンツの作成・収集」に示すように、デジタルコンテンツの作成時に併せて作成されることがよい。また、サムネイル/プレビューの活用のための権利処理についても、同様に、デジタルコンテンツの作成・収集時の留意点を参考にすることがよい。

(3) デジタルコンテンツの作成・収集

(デジタルコンテンツの作成時の留意点-品質について)

デジタルコンテンツの品質については、閲覧デバイス等の将来の高度化に備え、元のコンテンツの代替となるよう、コストとのバランスを考慮しつつ、可能な限り高品質なものを作成することがよい⁶。

また、高品質なものだけでなく、利用・提供のしやすさを優先して情報量を抑えたもの⁷や、発見を助けるためのもの（サムネイル/プレビュー）も併せて作成することがよい。サムネイル/プレビューの品質に関しては、利用する側が無理なく判別できるレベルのものが求められる⁸。さらには、Twitter 等の SNS で共有しやすいデータ容量の小さなものを用意することも考えられる。

なお、画像データ作成の際は、電子的に生まれるノイズや撮影時のピントによるボケ、天地等の確認を行う必要がある。また、色の管理については、印刷物による色校正ではなく、カラーチャートを入れることで撮影時の色の状態を残しておくことが求められる。

(デジタルコンテンツの作成時の留意点-契約や権利について)

デジタル化作業を外部に委託する際は、デジタル化成果物が自らの所有物となること、また、自ら自由に使えることに加え、第三者の活用も可能となるよう著作権の状態について、契約内容の確認を行うことが必要である。

⁶ 画像の場合、カラーチャートで色校され、メジャーを参照できるようにする。原資料・作品の大きさに対して 400dpi～600dpi でデジタル化し、ノイズの除去を行ったあと、非圧縮か可逆圧縮のフォーマットで作成することが望ましい。

⁷ 提供用画像としては、JPEG 等の非可逆圧縮フォーマットが考えられるが、自機関の回線環境等の状況に合わせて、適宜、解像度や画像圧縮率を検討することが求められる。

⁸ Europeana のサムネイル/プレビューの解像度の要件は最小で幅 400 pixel、国立国会図書館のデジタル化資料は 24 ビットフルカラーで幅 256 pixel(どちらも 2016 年末現在)。ネットワーク環境やデバイス全体の性能の向上に伴い、このレベルは随時見直されるべきものである。

展示会図録、報告書、レプリカ等の作成など、デジタルコンテンツの作成が主目的でない場合においても、その過程で様々なデジタルデータが作成される。市販の図録等の作成時に、元のコンテンツの著作権の権利処理を行っている場合は、デジタルアーカイブでの利用(データの公開や二次利用等)も併せて権利処理を行っておいた方がよい。

なお、美術・写真の著作物に関しては、美術館等の施設内において、作品の解説・紹介のために、作品の写真を小冊子に掲載することが現行の著作権法でも可能であるが、加えて、デジタル化したデータをタブレット等の電子機器で提供することが、著作権法の権利制限で可能となるよう、制度整備が進められている状況に留意が必要であろう⁹。

(デジタルコンテンツの収集時の留意点)

近年のスマートフォンの普及などにより、個人の所有するデジタル機器で簡単に高品質の写真、動画等を撮ることが可能となっている。例えば、地方のデジタルアーカイブにおいて、個人が撮影した祭り等の動画の記録を収集して自らのデジタルコンテンツとして増やしていくことも考えられる。こうした、個人が所有する写真・動画等の記録を収集する際、自らのサービスでの活用に加え第三者の活用も可能となるよう、包括的利用許諾を結ぶなど、権利関係の整備をしておくことがよい。その際、肖像権、プライバシー権等の諸権利にも留意が必要である。

(デジタルコンテンツの蓄積)

過去に作成したデジタルコンテンツについて、品質がやや劣るもの(白黒2値(モノクロ)や低解像度の画像など)であったとしても、活用に向けてそれらもデジタルコンテンツとして整備し、保存することがよい。

デジタルデータの保存に当たっては、画像・映像ファイル自体が持つ撮影時に自動的に記録された撮影日時・機器・解像度などの情報を削除しないよう注意する。削除する場合は、メタデータの一部として情報を保存する必要がある。これらのデータは、画像・映像データの表示のために再生機器を調整する際にも必要な情報である。

また、デジタルデータ作成時の情報が分かるよう、デジタル化の際のドキュメント等を残しておくことが求められる¹⁰。これら作成時の情報や作成フォーマット等については、デジタルコンテンツのメタデータに含めておくことがよい。

(参考情報:アーカイブ機関におけるデジタル化の法的根拠)

著作権法第31条に規定される「図書館等」に該当する博物館・美術館、図書館は、

⁹ 文化庁の「文化審議会著作権分科会法制・基本問題小委員会 中間まとめ」(平成29年2月)参照。

http://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/pdf/h2902_chukanmatome.pdf

¹⁰ ドキュメントのほか、作成されたデジタルデータがDVD等のパッケージで保存される場合、デジタルデータの存在が忘れられたりデータが失われたりするしないよう、そのパッケージ上に内容の記載を行い、保管場所を周知しておく必要がある。

所蔵資料・収蔵品のデジタル化に関し、現行著作権法上、保存のための複製を行うことが可能である。この点について、平成26年度の第41回文化審議会著作権分科会¹¹において、著作権法第31条第1項第2号に基づき、「図書館等」が保存のために「絶版等の理由により一般に入手することが困難な貴重な所蔵資料」をデジタル化することができるとの解釈が示されている。さらに、平成27年6月には、デジタル化できる主体の「図書館等」の範囲が拡大され¹²、博物館法の登録博物館及び博物館相当施設であって営利を目的としない法人により設置されたものが含まれることとなった。

(4) 長期アクセスの保証のために

(識別子の付与方法)

長期にわたり安定的にデータを保存・管理するため、アーカイブ機関は、個別の資料・作品や情報を判別・認識できる識別子(2(1)で示した管理番号)を付与することが求められる。

識別子の付与には、2種類の方法がある。組織内で決めたルールに基づき、一つは手作業で一意的な(重複しない)管理番号を付与する方法、もう一つは機械的に管理番号を付与する方法である。手作業で管理番号を付与する場合は、管理番号が重複しないように運用することが求められる。識別子(管理番号)は数字でもそれ以外でもよいが、国際的な流通を意識する場合、漢字等は使わない方がよりよい。

表形式により手作業でメタデータを整備している場合は、第一列目に、同一ファイル内で重複しない管理番号をつけるなどすればよい。ファイル内で一意であれば、「ファイル名」+「管理番号」を機関における識別子として利用できる。複数のファイルを複数のフォルダごとに管理している場合には、それらを「/(スラッシュ)」でつないで「フォルダ名/ファイル名/管理番号」などとすればよい。

機械的に管理番号を付与する場合は、一意に特定するための管理番号を付与するルールとなるよう、プログラムなどを調節する必要がある。データベースシステムを用いてメタデータ等を整備している場合は、一意の管理番号が自動で付与されるように設定することがよい。

手動にせよ機械的な付与にせよ、一度付与した識別子(管理番号)は、原則として変更しないようにしなければならない。同じ番号を別のデータに付与するような使い回しをしてはならないし、機械的な付与においては、削除した場合に自動で新たな番号が振り直されることのないよう気をつける必要がある。

なお、メタデータを新規に整備する場合、自館の所蔵品目録や展示会図録等から全作品の一括付与を仮に行う方法なども考えられる。その識別子を用いて自ら保有する作品情報や解説等のデータを整備・公開するなどの活用が可能になる。

¹¹ <http://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/bunkakai/41/index.html>

¹² 平成27年文化庁告示第30号

(識別子の種類)

識別子の付与は、データを共有する際の相互運用性の担保においても重要な取組である。URI(Uniform Resource Identifier)の付与については「3(5)データ共有の方法」を参照のこと。

このほか、デジタルコンテンツの識別子として、国際標準規格であるDOI(Digital Object Identifier)がある。アーカイブ機関が個別のコンテンツごとに URL を指定してアクセス可能なウェブページを提供している場合は、国際的な流通促進を考慮して、DOI を付与することも考えられる。DOI は永続的な識別子であるため、DOI のメタデータのメンテナンスが必要であるが、URL が変更されても長期アクセスが保証される点にメリットがある。DOI の付与には、我が国で唯一の DOI 付与機関であるジャパンリンクセンター(JaLC)¹³の会員になる必要がある。

機関を特定するための識別子としては、国際標準規格である ISIL (International standard identifier for libraries and related organizations)があり、我が国では、国立国会図書館が付与・管理を行っている¹⁴。つなぎ役が機関コードの管理を行う場合に ISIL の活用が考えられる。

(システム持続可能性・安定性の確保)

アーカイブ機関は、デジタルアーカイブの持続的なアクセスの維持に関して、一定の配慮を行う必要がある。プラットフォームやサイトについて、安定性を確保するため、システム更新や、組織の統廃合によってデータが失われないよう、データ作成時・システム構築時から、画像・映像等のデジタルコンテンツやサムネイル/プレビューは、メタデータも含め、特定の機器(システム、メディア)等に依存しないデータ形式とし、データ移行性を確保することが重要である。ストレージ装置や各種デバイス、システム自体は数年ごとのリプレースが必須であり、そのための経費、さらにメンテナンスに従事する人員の確保も含め、運用面のコストをデジタルアーカイブ構築時にある程度見込んでおくことが求められる。

(長期アクセスの保証)

アーカイブ機関は、保有するコンテンツの保存に加え、メタデータやサムネイル/プレビューを含むデジタル情報資源の保存にも取り組むことが求められる。

2(1)で既述したメタデータ項目については、それぞれの項目の意味が将来的にも把握できるようにドキュメントを整備し、維持する必要がある。

デジタルコンテンツの保存のためには、長期的な利用可能性を考慮し、特定製品等に

¹³ <https://japanlinkcenter.org/>

¹⁴ <http://www.ndl.go.jp/jp/library/isil/>

なお、現在 ISIL は図書館中心の登録であるため、他の分野に属する機関の登録への広がりが望まれる。

依存せず、仕様等が公開され、かつ広く普及している(国際標準等で定められた)フォーマットを採用する必要がある。また、2(3)で既述したとおり、デジタル化の際のドキュメント等の記録を残しておくことも求められる。

さらに、メタデータ、サムネイル/プレビュー、デジタルコンテンツなど関連する全てのデジタル情報資源にいえることであるが、3章で述べるようにデータの共有を促進することにより、複数箇所にデータが存在する状況を作り出すことも、保存の有効な方策である。災害や大規模なシステム障害等への対応可能性が高まることから、データの分散化・複数化を進めることが求められる。

なお、長期的な保存を考えたアーカイブの仕組みに関する標準については、欧米ではOAIS参照モデル(参考資料「確認すべき標準・ガイドライン等」参照)を基にした各種ガイドライン等が整備されており、我が国でも留意することが望ましい。

3. データを共有するに当たって

アーカイブ機関は、自ら作成・保有するデジタル情報資源について、二次利用の条件も含めてどのように公開し提供していくのかについて、著作権等に配慮したうえで自ら決定して共有・発信することができる。ただし、公的機関が著作権を保有するもの又は公的助成により作成されたデータに関しては、できる限り広く活用可能な形で共有・発信していくことが求められる。

アーカイブ機関は、公開範囲と公開内容を決定するに当たっては、本章に示す点に留意して共有・発信していくことがよい。

なお、ここでいう共有の対象となるデータは、デジタルアーカイブで扱うデータ全般とし、デジタルコンテンツはもちろん、メタデータやサムネイル/プレビューを含む。

(1) 公開ポリシーの考え方

アーカイブ機関は、自らが作成・保有するデジタル情報資源について、それぞれ種類ごとに公開範囲を決めるとともに、第三者が二次利用する場合の条件についても決めることができるし、決める必要がある。

例えば、メタデータ、サムネイル/プレビュー、デジタルコンテンツなどの種別に応じて、それぞれ異なる公開範囲を決めてもよいし、さらにデータが持つテーマや品質などによって公開範囲を分けてもよい。

公開範囲としては、以下の選択肢がありうる。

- ① 自らの業務用限り
- ② 自らの施設内でのサービス利用
- ③ 自らのホームページ上での発信
- ④ 同じ分野・地域のコミュニティネットワーク内での共有
- ⑤ 国の分野横断統合ポータルでの検索結果の表示
- ⑥ 第三者に二次利用を可能とする(自ら発信する③に限らず④や⑤を経由して公開する場合も含む。)

①の場合はコンテンツ管理のためであり、②の場合は来館者への案内や展示会の解説、更にはデジタルデバイスを用いた施設内での体験学習など、自らのサービスの充実に活用できる。

③の場合は、遠隔地にいる非来館者に対して、自らの存在を確認してもらうことができるうえ、デジタルコンテンツを提供する場合は、自らの魅力を示すことで存在価値を高めることができる。

④の場合は、同じ分野・地域のネットワークの中で、データを共有することで、来館者に

対する、更に充実したサービスが展開できる。例えば、同時に同じテーマの展示会やイベントを複数の施設で開催するなどの取組が考えられる。

⑤の場合は、③又は④による提供方法の拡充につながるもので、国の分野横断統合ポータルにおける検索を可能とすることである。

⑥の二次利用については、③、④及び⑤の共有・発信の検討に当たって、併せて検討されるべきものである。例えば、教育目的は可だが営利は不可、誰でも自由な利用が可など、二次利用条件の選択肢は複数ある。その詳細は、次項以降に示すが、⑥に対応することは、デジタルアーカイブ社会の推進にとって、最も望ましく重要な取組である。

(2) 二次利用条件の表示方法

(利用条件の基本的考え方)

利用条件付与の検討においては、先ず当該データの権利の状態を確認する必要がある。アーカイブ機関のみが権利を有するデータの場合は、自らがその二次利用の条件を設定することができる。第三者が権利を部分的にせよ有し、かつ包括的な許諾などがなされていない場合は、どのような条件の利用とするかについては、当該第三者と協議し、合意と許諾を得る必要がある。

なお、二次元の作品を正面から撮影した場合や、三次元の作品であっても三面図的に記録した場合は、撮影者やデータ作成者の著作権は原則として認められない。ただし、特定の角度、照明等により撮影者の芸術表現として撮影された写真等、撮影者の創作性が認められる場合は、撮影者の著作権の権利が発生する場合があることについて、注意が必要である。

著作権保護期間が満了しているコンテンツを撮影したもので、かつ写真撮影者にも著作権が発生しない画像データや、創作性のないメタデータなど、著作権法による保護対象とならないデータについては、原則として権利の問題は発生せず、営利・非営利を問わず誰でも自由に利用可能とされている。ただし、プライバシー権や著作者人格権¹⁵等の配慮が必要な場合があることに留意しなければならない。

また、著作権法による保護の対象とならないデータであっても、そのデータの活用においては、データ提供元の各アーカイブ機関やデータ作成者等の貢献を社会的に認知してもらうため、またデータの信頼性を担保するため、活用者に対して、二次利用に際し出典や所蔵館等の表記を正確な形で行ってもらえるよう、望ましい表記事項等のお願いをアーカイブ機関がウェブ上に掲載することは考えられる。本ガイドラインにおいても、「4(1)データの活用における留意点」及び「4(4)活用の結果できた成果物の還元」に、活用者に対して出典等の記載を行うのが望ましいとしている。

¹⁵ 我が国の著作権法においては、著作者の死後においても、その行為の性質及び程度、社会的性情の変動その他によりその行為が当該著作者の意を害しないと認められる場合を除き、著作者人格権の侵害となるべき行為をしてはならないと規定されている。

(利用条件表示の種類)

著作権の保護対象であるデータの利用条件表示は、一定の自由利用を認める場合、原則としてクリエイティブ・コモンズ(CC)¹⁶を使うことがよい。デジタルアーカイブのデータの国際的な活用促進のためには、世界的に主流となっている条件と共通化を図る必要があるためである。クリエイティブ・コモンズは、我が国を含めた世界各国のオープンデータ(オープンガバメント)の取組においても、広く活用されている。

具体的には、クリエイティブ・コモンズ CC0(CC0)、6種類のクリエイティブ・コモンズ・ライセンス(CC BY、CC BY-SA、CC BY-ND、CC BY-NC、CC BY-NC-SA、CC BY-NC-ND)によって利用条件を表示することがよい。著作権保護期間が満了しているなど著作権による制限がないデータは、パブリック・ドメイン・マーク(PDM)などを利用してそのことを明示することがよい。それぞれの意味は、補足資料「利用条件表示一覧」を参照のこと。

このほか、我が国には、「政府標準利用規約(第2.0版)」があるが、これはCC BY4.0と互換性があるとされている。国際的な流通を容易にするには、CC BY4.0の表記も示すことが望ましい。さらに、上述のライセンス等を補完する情報として、著作権者不明等の場合の裁定制度を利用している場合は、そのことを示す情報を併せて提供する必要がある。また、我が国の法律上著作権の保護期間が満了していることを併せて提供することが望ましい。

また、利用に当たって何らかの登録や対価を求める場合など、自由利用以外の場合については、デジタルアーカイブで提供されているデータがどのような手続で利用できるのかについて、メタデータや提供ページ等で明示的に示すことがよい。

(利用条件の表示方法)

利用条件の表示においては、ウェブページに利用条件のマークを示して目で見て分かるようにするだけでなく、機械的に解読可能な形式でも提供できるようにする必要がある。

特に、コンテンツの利用条件表示においては、コンテンツのデータファイル自体に記述するだけでなく、メタデータの項目にも利用条件の情報を保持することがよい。このためには、「2(1)メタデータの整備」で用意したメタデータに、コンテンツごとの利用条件を追加する必要がある。各コンテンツに一つ一つ利用条件を追加するのが困難な場合は、一括して処理できるよう、権利等の状態が同じコンテンツのメタデータは、まとめて管理しておくことよい。

また、利用条件や権利の内容についての情報を提供するページへのリンクもあることがよい。その際、ページでの説明が日本語のみの表記ではなく、多言語(英語等)での表記も用意されていることが望ましい。

クリエイティブ・コモンズ・ライセンスやCC0を適用する対象は、本来、著作権の保護対象であるため、著作権保護期間が満了している所蔵作品のデジタル化データにそれらを適用する場合は、その所蔵作品の作者がアーカイブ機関自身であるなどの誤解を招くことがないように、アーカイブ機関自らがそのデータに関してどのような権利を保有しているか

¹⁶ 本ガイドラインでは、現時点の最新版であるCC4.0の採用を推奨する。

(例えば、創作性のある写真の撮影者としての著作権を有する等)を明記することが望ましい。

(3) 望ましい利用条件(オープン化の推進)

デジタルアーカイブで提供するデータにどういった利用条件を付すかを判断するに当たっては、活用が最大限に行われるよう、可能な限りオープン化し、共有する各種データが自由に二次利用できることが望まれる。本ガイドラインでは、利用条件として、PDM、CC0、CC BY を表示することを推奨する。

以下、特にその中でも、メタデータ、サムネイル/プレビュー、デジタルコンテンツごとに望ましい条件を示す。

(メタデータの望ましい利用条件)

何をオープンにし、何をオープンにしないかは、アーカイブ機関が自ら判断するものではあるが、メタデータの利用条件の設定は、各コンテンツの発見可能性に直結する、国のデジタルアーカイブ振興における中核的課題である。特に、国際的なメタデータの流通・活用を進めるには、たとえメタデータの項目等を国際標準にそろえても、国際的に普及している基準と同等の利用条件でなければ、グローバルな流通は難しい。例えば、国の分野横断統合ポータルから、EuropeanaやDPLAといった海外の主要ポータルと連携していく場合、かれらと同等の利用条件でなければ、独自の利用条件を理解するための手間がメタデータ活用の障壁となり、国際的なメタデータ流通から後れを取ることになる。

このため、公的機関のメタデータ又は公的助成により作成されたメタデータについては、世界で流通しているメタデータと利用条件のレベルをそろえ、CC0 を採用することが求められる。第三者が保有する著作権等の権利を根拠に CC0 を設定できない場合は、その制限される部分(項目)とその理由、並びに利用条件を明示することが必要である。

民間の機関が提供するメタデータに関しては、同様の考え方から、少なくとも国の分野横断統合ポータルや同じ分野・地域のコミュニティなど他のデジタルアーカイブと連携して共有するためのメタデータについては、CC0を採用することが望ましい。

(サムネイル/プレビューの望ましい利用条件)

オープンデータ(オープンガバメント)やオープンサイエンスに係る関係府省の施策との一貫性の観点も踏まえ、公的機関のものや公的助成を受けて作成されたサムネイル/プレビューについては、第三者が保有する著作権等の権利が問題とならない限りは、原則として CC0 又は(政府標準利用規約と互換性のある)CC BY を適用することが求められる。公的機関は、第三者の著作権等を根拠に CC0 又は CC BY が設定できない場合には、その制限される理由及び利用条件を明示することが必要である。

著作権保護期間が満了しているなど著作権による制限がないものは、パブリック・ドメイン・マーク(PDM)などを利用してそのことを明示することがよい。元のコンテンツの権利が

パブリック・ドメインのサムネイル/プレビューは、著作権の状態を確認しつつ、可能な限り、パブリック・ドメイン・マーク(PDM)などを表示し、自由に利用可能であることを明示することがよい。

サムネイル/プレビューの元のコンテンツが著作権保護期間内の資料・作品である場合は、権利の問題が発生する。デジタルアーカイブの検索結果としてサムネイル画像等をインターネット送信することは、現行の著作権法では、著作権者の許諾が必要である¹⁷。許諾を得るに当たっては、所蔵館での利用に限った内容とせず、連携先のデジタルアーカイブや外部のポータルサイトでの検索結果としても表示できるよう、更には、活用者がダウンロードして利用できるよう、CC0 又は CC BY 等のオープンな条件での提供を可能とする内容の許諾を得るようにすることが望ましい。

(デジタルコンテンツの望ましい利用条件)

作成されたデジタルコンテンツがより多くアクセスされ、SNS 等での言及が増えたり、様々な形で活用されれば、コンテンツの注目度が高まり、その社会的な価値も結果として高まる。デジタル化の元となった実物そのものについても同様である。アーカイブ機関は、自らのみが権利を有している限りは、自ら望ましいと考える利用条件を付すことができる。無償での再利用に問題がないものについては、インターネットで公開し、CC BY 等のオープンな条件で提供することが望ましい。

著作権保護期間が満了しているなど著作権による制限がないものは、パブリック・ドメイン・マーク(PDM)などを用いて、自由な利用が可能であることを明示しておくことがよい。

公的機関の、又は公的助成により作成されたデジタルコンテンツについては、無償での再利用に問題がないものについては、原則として CC0 又は(政府標準利用規約と互換性のある)CC BY を適用することが求められる。

なお、公的機関・公的助成によるデジタルコンテンツであっても、収益事業等、事業の運営に甚大な影響がある場合は、一定程度の配慮をした利用条件を付すことも妨げない。

表2: 公的機関のもの又は公的助成により生成されたデータの二次利用条件

データ種別	自らが著作権を保有するものの二次利用条件
メタデータ	CC0
サムネイル/プレビュー	CC0、CC BY、(PDM)
デジタルコンテンツ	CC0、CC BY、(PDM)

¹⁷ ただし、文化庁の文化審議会著作権分科会法制・基本問題小委員会において、美術・写真の著作物に関しては、原作品を適法に展示する者が、展示作品に関する情報を広く提供するためにサムネイル画像をインターネット等で公開すること、美術館等の施設内において展示作品の解説・紹介のためにデジタル化データを作成し電子機器に掲載・上映することについて、権利制限の対象となるよう検討が進められているところである。(注記 9 参照)

(4) 利用条件表示の検討に当たっての留意点

単なる事実や数値を記述しただけのデータであれば、著作物性が認められないため、著作権保護の対象にはならない。編集著作物やデータベースの著作物と認められる場合であっても、素材となるデータに著作物性がない場合は、データそのものを利用することは著作権法の観点からは制限されないことに留意する(「二次利用の促進のための府省のデータ公開に関する基本的考え方(ガイドライン)」参照)。一方で、著作物性の有無について活用者が逐一厳密に判断することは困難であり、また、著作物性の判断基準は国によっても異なることから、誰もがグローバルに確実に自由利用可能であることを担保するため、本ガイドラインでは、メタデータに関して、確認的な意味を含めてCC0の適用が望ましいとしている。

CC0とは、全ての著作権等の権利を放棄することを意味する。これは、著作権に基づいて訴訟を起こす権利、逸失利益等が出て損害賠償を求める不法行為に基づき訴訟を起こす権利も含めて放棄し、著作者人格権など放棄できない権利については行使しないことを約束するといったことを意味する。

著作物性のあるデータであっても、著作権保護期間が満了している場合は、できる限り、PDMを表示するなどして、自由な二次利用が可能であることを明示することが求められる。ただし、現時点では、我が国で著作権保護期間が満了していても、欧米等では保護期間内である作品も多く存在するため、PDMを表示する際は、少なくとも著作者の死後70年以上経過するなどして、海外においても著作権保護期間が満了していると考えられるものに対して付与する必要があることに留意しなければならない。

また、クリエイティブ・コモンズ・ライセンス等の著作権ライセンスは、著作権者が自ら保有する権利の利用許諾を行うものであるため、著作権保護期間が満了しているコンテンツや、著作物としての要件を満たさない創作性のないデータなどに適用したとしても、その二次利用に著作権法による保護を超えた追加的な制約を課すものではないことに留意が必要である。

なお、我が国の著作権法上で保護期間満了のものや著作権者不明等の場合の裁定制度に基づく利用については、現状は統一したマーク等の利用条件表示の手段が存在せず、海外の活用者にも理解が容易な、機械による判別が可能な仕組みの用意が難しい点は今後の課題である。

アーカイブ機関で利用条件を検討するに当たっては、著作権のほか、肖像権、パブリシティ権、プライバシー権¹⁸等の諸権利にも留意が必要である。

また、利用条件の検討において、セキュリティポリシーが問題となって公開が妨げられることがないよう、所蔵資料・収蔵品の目録等の情報については、広く共有されることが

¹⁸ 公文書館が個人情報等を一定期間公開しない措置を取る際の基準については、参考資料「確認すべき標準・ガイドライン等」にあげている「独立行政法人国立公文書館における公文書管理法に基づく利用請求に対する処分に係る審査基準」の別添の情報が参考になる。

望ましいという観点から検討を行い、単なる自らの組織のデータ保全を行うという観点とは切り分けた判断を行うことが求められる。

(5) データ共有の方法

アーカイブ機関がメタデータやデジタルコンテンツ等を他機関と共有したり、広く発信したりする場合、テキストデータや画像データ等のファイルをそのままウェブに掲載する方法もあるが、一般的にはユーザが手軽に検索できるデータベースの形式で提供することが望ましい。さらに、組織内にとどまらず、分野・地域のコミュニティネットワーク内での共有や国の分野横断統合ポータルでの利用までを視野に入れた場合は、機械可読性を確保し、サーバ間でデータのやりとりが可能となるよう、データベースにコンピュータ向けインターフェース(アプリケーション・プログラミング・インターフェース、API)を設けることが望ましい。

自ら発信する場合のパッケージソフトウェアの選択や、発信することが困難な場合にどの分野・地域のコミュニティのプラットフォームと連携するかについては、こうしたデータのやり取りを容易にする機能が備わっているかどうか、大きな判断基準となる。また、独自に開発等を行う場合には、こうした機能について、仕様に明記することが求められる。

以下、パッケージソフトウェアやプラットフォームが備えているべき、具体的な機能について述べる。

(国際的にデータを共有するための多言語化対応)

データの活用を最大限有効なものとし、国際的な流通を促進するには、インターフェースやヘルプ類の多言語化、特に英語版の提供が求められる。少なくとも資料・作品のタイトル等の識別に重要な項目だけでも、英語又はローマ字表記が用意できるとよい。

また、分野や地域のコミュニティの中で「つなぎ役」を務める機関の場合は、国の分野横断統合ポータルや海外の主要なポータルとの連携を視野に入れたデータ共有の仕組みを考えることが望まれる。

(識別子となるURIの付与)

アーカイブ機関は URI(Uniform Resource Identifier)を識別子としてメタデータに付与し、提供することが求められる。URI の付与においては、既存の管理番号などを用いて一括生成するのが合理的である。特に「2(3)長期アクセスの保証のために」で付した識別子がURIに含まれるような形で設定すると、URIとメタデータの識別子の対応関係が容易に判別できる。

なお、URIの付与をアーカイブ機関が自ら行えなくとも、メタデータの管理ファイルをウェブで公開していれば、デジタルアーカイブの活用者が自らの利用の利便のために付与してくれる場合がある。ファイル単位で管理番号を唯一にしていれば、活用者は、ファイル

のURLと管理番号を組み合わせることでURIを構成できる。ただし、その場合、メタデータのファイルをウェブ上の安定的な(URLが変更されない)ところに置かなければならない。また、ファイルの形式は、参考資料「活用できる表形式のデータとは？」の要件を満たすことが求められる。

(メタデータの共有方法)

メタデータの提供においては、以下に示す方法のいずれかが用意されていることが求められる。特に分野や地域のコミュニティの中で「つなぎ役」を務める機関の場合は、国の分野横断統合ポータルや海外の主要なポータルとの連携を視野に入れたデータ共有の仕組みがあることが求められる。具体的には、以下に述べるOAI-PMH等のAPIを備えている必要がある。

また、コンテンツも併せて提供する場合には、メタデータにコンテンツのURLも自動的に組み込む形で提供する機能が備わっていることが望ましい。

① OAI-PMH

OAI-PMH は、機械的にシステム間でメタデータを収集・差分更新する API を実現するために現在最も普及している標準である。パッケージソフトウェアを選択したり、システムの導入を行う際には、OAI-PMH に対応したものを採用したり、導入要件として仕様書に記載することで、多くの機関やサービスとの連携が期待できる。

OAI-PMH は、メタデータを収集するためのプロトコルを使用して複数のアーカイブからメタデータを収集し、更に収集した側で加工処理してサービスを提供するものである。条件を細かく設定して一括してデータを取得できること、差分収集ができることがメリットとしてある。

分野・地域のコミュニティの「つなぎ役」がデータ連携方法として OAI-PMH 方式の API を採用し、メタデータを収集(ハーベスト)している場合は、参加するアーカイブ機関側の使用するシステムについても、同じ方式に対応していることが望ましい。

② Linked Data

Linked Data は、ウェブ上で多様な種類のデータを結びつけて共有するための技術である。アーカイブ機関が連携するためのパッケージソフトウェアを選択したり、仕様書を書く際には、自らのデータをより活用してもらうため、Linked Data に対応したものを選択することが望ましい。今後、国や分野・地域コミュニティごとのプラットフォームを整備する際にも、Linked Data への対応は重要な位置を占めるものとなる。

Linked Data は、対象の識別子として与えた一つの URI をリンク先として要求すると、その対象についてのメタデータを返戻してくれる仕組みである。例えば、あるアーカイブ機関の所蔵作品に URI が付与され、そこからメタデータを取得できれば、そのメタデータの中から作者の情報を抽出し、別の機関が提供する当該作者に関するデータをリンクさせることで、より詳細なデータを取り込んで活用したり、当該作者と知人関係にある人物のリストを作成したり、さらには、所蔵作品に関する複数の言語で出版されている翻訳物のリ

ストや関連テーマの作品リストを作成できたりする¹⁹。フォーマットやデータ構造の異同にかかわらず、自らが発信するデータと他機関の関連するデータとを結びつけることで様々な活用が可能になるメリットがある。

Linked Data に対応するためには、アーカイブ機関は、前述の URI による識別だけでなくメタデータを機械的に取得できるようにする必要がある。このため、URI を簡便に付与できる機能に加え、その URI に対して何らかのアクセス機能があるパッケージを選択するか、又は仕様書にこれらの機能の対応を記載することが望ましい。

③ その他 API

Linked Data に対応する技術として、SPARQL Endpoint を使った検索ができる API の機能を備えているとなおよい。SPARQL は、RDF(Resource Description Framework)化して蓄積したデータを検索するもので、検索結果として表又は RDF データが返戻されるものである。アーカイブの活用者は、構造化されたデータをまとめて検索結果として扱うことができるため、②の Linked Data での提供に加えて、より活用の幅が広がることになる。

このほか、横断検索を実現するための API として、SRU、SRW、OpenSearch、OpenURL などの仕組みがある。これらの API は、可能であれば、①の OAI-PMH や②の Linked Data に加えて、いくつか用意されているとなおよい。デジタルアーカイブを利活用する側が様々な方法でメタデータを活用できるよう、複数のメタデータ連携の仕組みが備わっているソフトウェアやプラットフォームを選択することが望ましい。

④ オープンなデータをダウンロード可能にしておく方法

パッケージソフトウェアの導入や、プラットフォームへの参加が困難な場合、①～③の機能の独自整備は費用的にやや割高であり、技術的な知識を必要とするため、各アーカイブ機関で対応することは難しい面がある。

こうした場合には、表形式のデータ(CSV、TSV 等)をウェブ(ホームページ)で提供することも有用な連携手段になりうる。また、SPARQL Endpoint を実装せずとも、RDF データをファイルとしてウェブ上の安定した場所に置き、データ取得を可能としておく連携形式も考えられる。

ただし、こうした形式でのデータでの連携は、分野・地域のコミュニティのつなぎ役など、データ取得を行う側の機関やサービスで人的・費用的コストがかかるため、取得する相手の状況に留意が必要である。なお、連携先で有用となる表形式データとするために最低限求められる条件は、参考資料「活用できる表形式のデータとは？」を参照のこと。

(サムネイル/プレビューの共有方法)

サムネイル/プレビューについては、コンテンツの理解を助けるために検索結果で提供

¹⁹ 具体的な活用事例については、国立国会図書館ホームページの「使う・つなげる：国立国会図書館の LOD でつながる」を参照されたい。

http://www.ndl.go.jp/jp/aboutus/standards/lod_usecase.html

される仕組みを備えていることが求められる。コンテンツへの有効なナビゲーションとなるだけでなく、高精細画像をメタデータとともに大量にリアルタイムで提供する必要がなくなるため、システム面でのコスト負担が軽減される。

サムネイル/プレビューからは、デジタルコンテンツへのリンクがあるとなおよい。

サムネイル/プレビューの共有・発信は、デジタルアーカイブを利用する側にとってもメリットが大きい。検索結果での表示のほか、サムネイル/プレビューがメタデータとセットで自由な二次利用が可能になると、より様々な活用の幅が広がる。そのためには、サムネイル/プレビューのURLがメタデータ項目の一部として自動的に提供される機能があるシステムを導入することが望ましい。

(デジタルコンテンツの共有方法)

相互運用性の確保のため、デジタルコンテンツを提供する際は、広く標準的に使われており、特定の製品に依存しないフォーマットに対応しているシステムを導入することが求められる。

また、異なるデジタルアーカイブ間において、デジタルコンテンツ同士を一緒に利用できるようにする仕組みが用意されているとよい。例えば、画像データの共有の仕組みの一つに、海外の主要なアーカイブ機関が採用している国際的な標準 IIIF (International Image Interoperability Framework) がある。IIIF は、利用する側だけでなく、提供する側にとっても、どこで見られたのかを確認できるため、メリットが大きい。現在は IIIF に対応したパッケージソフトウェアは少ないが、今後の対応可能性を選択する際には考慮することが望まれる。

この他、Twitter等のSNSに簡単に投稿できる仕組みが用意されているとよい。具体的には、画像添付ボタンや、ツイート用の容量の少ない画像の提供機能などが考えられる。

4. データを活用するに当たって

博物館・美術館、図書館、文書館、大学・研究機関、企業、市民団体、官公庁・地方自治体等といった「アーカイブ機関」が自らのために活用するほか、一般ユーザ、IT 技術者、クリエイターなど、様々な機関・団体・個人がデジタルアーカイブで提供される各種データの活用者になりうる。オープンなデータを使って、電子展示会の開催、観光用 VR のアプリ提供、教育目的でのコンテンツ利用、防災ポータル構築、人工知能(AI)の学習用データ及び研究対象のデータとしての活用など、様々な目的での活用方法が考えられる。こうしたデータの発見可能性を高めるのが、国の分野横断統合ポータルの役割であるといえる。

本章では、オープンになっているデジタルアーカイブのデジタル情報資源を最大限に活用し、かつ、データを提供する側のメリットにもなるよう、データの活用者が何に留意して利活用すべきか、何をすることが求められているかを示す。

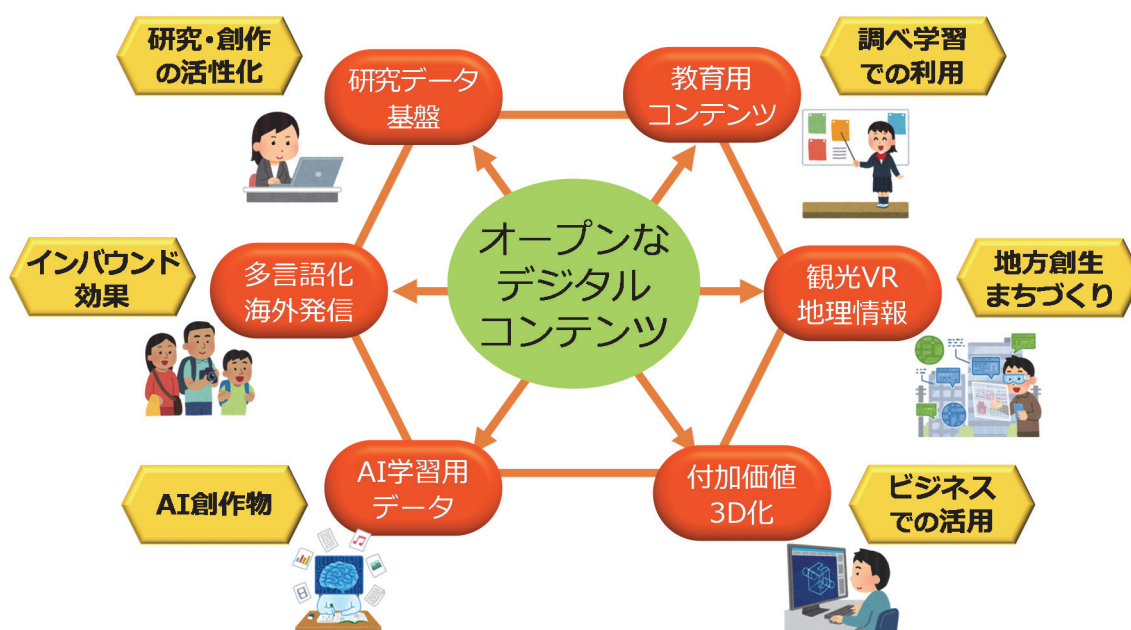


図 4 デジタルアーカイブ社会における活用イメージ(例)

(1) データの活用における留意点

(活用における基本的な考え方)

アーカイブ機関が広く提供しているデータに関し、活用者は、そのデータ提供者が示す二次利用の条件等を踏まえた適切な活用が求められる。

アーカイブのデータを利活用するに当たっては、その分野のコンテンツの理解に努め、できる限り、コンテンツ自体の価値をさらに高め、提供者にとってもメリットにつながる形で活用することが求められる。

(ライセンス等における留意点)

著作権の保護対象であるデジタルアーカイブのデータ活用に当たっては、適用されているライセンスや利用条件をよく確認し、順守しなければならない。主要なライセンスについては、補足資料「利用条件表示一覧」を参照のこと。

PDMなど著作権保護期間が満了していることが明示されているデータや、完全に権利が放棄されたデータであることを示すCC0が適用されたデータであったとしても、著作者人格権やプライバシー権等への配慮に加え、データ提供者や作成者等の貢献を社会的に認知してもらうことへの配慮、データの信頼性の担保といったことにも留意することが必要である。具体的には、可能な限り、作者名等の出典に加え、提供元であるデータ提供者等のクレジットや、元データのURIを示すことが望ましい。

また、ライセンスによって自由利用が明示的に認められている場合であっても、作品の中に写り込んだ第三者の肖像権やプライバシー権等については、別途確認が必要な場合があることに留意する必要がある。

(2) 付加価値情報の付与

(新たな情報の追加)

活用者は、デジタルアーカイブで提供されている様々なデータに対して、付加価値となる情報を追加して利用することが求められる。これにより、(1章で示すところの)活用領域にある個々のデジタルアーカイブのデータが充実したものとなる。活用者による情報の追加は、データ提供者にとってもコンテンツに対する新たな価値付けによるメリットが生まれるため、オープンに発信するインセンティブにもつながる。

付加する情報には、活用者自らが知っている情報を追加することもあれば、海外発信の強化を助けるための、多言語化(英語等)やローマ字表記の追加といったことも考えられる。今後は、APIと自動翻訳を用いた、他言語による検索サービスの提供なども考えられる。

また、Linked Dataによって活用領域にあるウェブ上の様々なデータをつなぐことで、異なる機関が提供するデータを関連付けることも考えられる。例えば、ゴジラに関する映画や出版物、その出版物を所蔵している図書館といったデータをつなぐことで、ゴジラをキ

一として様々な情報を追加していくことができる。データを提供するアーカイブ機関がURIを付与していない場合でも、安定したURLさえ提供できていれば、活用者がウェブ上にあるデータを使ってLinked Data化を進めることが可能である。

ポータルやアプリの作成においては、活用者が自ら知っている情報やすでに活用領域にあるデータをつなげるだけでなく、そのポータルやアプリを利用する人々が、知っている情報を自由に追加できるような、利用者参加型の仕組みをサイト等に用意する方法も考えられる。例えば、せんだいメディアテークの「どこコレ？」²⁰では、昔の写真や記録映像について、それがいつ頃で場所がどこかについて、参加者が自分の知っている情報を記録していくといった試みが行われている。

(追加に当たっての留意点)

活用者が新たな情報を追加した際には、元のデータに何の情報を追加したか分かるような形で、活用したデータを提供することが求められる。また、追加したデータについても、オープンな(自由な二次利用が可能な)条件で公開することで、さらなる情報の追加が連鎖していくことが期待される。

(3) 情報間の関連付け

(情報をつなげるために必要な作業)

活用者は、多種多様なデータの中から、関連する情報をつなげてメタデータをより豊かにしていくことが望まれる。その際、オープンな利用が可能なメタデータについて、活用者は、

- ・地理情報
- ・時間情報
- ・人物情報

などの分野を横断して共通する情報を用いて、異なるアーカイブ機関間で提供されているメタデータを関連付けていくことが考えられる。これにより、機関・分野を超えたコンテンツ間の関係性を創りだすことができ、地理情報や時間情報は、地図や年表などに活用できる。例えば、地図上に文化財のデジタルコンテンツや関連するデジタル化資料を示すことによって、観光客に役立つアプリの作成や、小説や漫画に出てくるスポットの名称と作品名を示すアプリ²¹も作成できる。また、時間順に作成された美術作品を並べてそれぞれの所蔵館を示すといった仕組みも考えられる。

データを関連付ける際、「3(5)データ共有の方法」で既述した Linked Data を用いる方法がある。あるアーカイブ機関が提供するメタデータの中に、他から提供されている人物や場所を示した URI のリンクを追加するという方法である。例えば、人物や地名の場合は、

²⁰ <http://www.smt.jp/projects/doko/>

²¹ http://www.ndl.go.jp/jp/data/bib_newsletter/2015_3/article_01.html

国立国会図書館の Web NDL Authorities²²の URI を用いることが考えられる。

(情報をよりつなげやすくするために)

より効果的な情報の関連付けのためには、分野のコミュニティ内で当該分野に特化した専門的な人名・地名を示す情報を集約させて URI を付与し、Wikidata、DBpedia などに識別リンクを提供するなどして URI を増やして活用しやすい環境を整備することが求められる。

情報の交換・共有をより有効なものとするためには、分野のコミュニティ内でよく用いられる標準的な用語を統制するため、辞書・典拠・シソーラスといった管理を行うことが求められる。例えば、日本美術の作家に関する人名の表記について、複数の表記方法をまとめるといった作業である。また、用語の統制に加えて、同じ分野でのメタデータフォーマットの標準化が行われることも求められる。

これらの作業は、分野や地域のコミュニティにおいて、つなぎ役が中心となって、分野横断的なメタデータ要素と各分野固有のメタデータ要素を整理し、連携するために共有すべき要素を明確化することが望ましい。こうしたメタデータの標準化や用語の管理は、活用するために必要なだけでなく、それぞれのアーカイブ機関のメタデータ整備に役立つ、業務の効率化につながるものでもある。

(4) 活用の結果できた成果物の還元

(保存・共有領域への還元のために)

活用者は、同時に、データの提供者にもなりうる。(1章で示すところの)保存・共有領域で行うべきこと、特に以下の点について留意することが求められる。

- 二次利用条件の表示とオープン化の推進等(3(2)~(4)参照)
 - ・可能な範囲でオープンな利用条件(CC0、CC BY)での提供
- データ共有の方法(3(5)参照)
 - ・Linked Dataによる活用の広がり確保(URLの付与等)
 - ・デジタルコンテンツの相互運用性の確保(画像の場合はIIIF対応)
- 長期アクセスの保証のために(2(4)参照)
 - ・識別子の付与
 - ・システム持続可能性・安定性の確保
 - ・長期アクセスの保証

また、活用者は、更なる活用を促進できるよう、デジタルアーカイブのデータを使った成果について、Twitter等のSNSやウィキペディアなどに積極的に発信していくことも望ましい。

²² <https://id.ndl.go.jp/auth/ndla>

(データ提供者への還元のために)

活用者が追加した付加価値情報や関連付けした情報などは、データの提供者であるアーカイブ機関やその分野・地域のコミュニティに(つなぎ役を経由するなどして)情報がフィードバックされることが望ましい。これにより、充実したデータがより適切な形で流通することになる。活用者は、それを意識した形でデータを活用することが求められる。

活用者は、データをオープンに提供している機関のため、どのような活用がなされたか把握できるよう、著作権の保護対象にならないデータやCC0が適用されているデータを利用する場合であっても、4(1)に既述したとおり、可能な範囲で出典や所蔵館などの情報を明示する、リンクを貼るといったことを行うのが望ましい。

利用したデータの著作権保護期間が満了している場合は、それを利用して提供するデータについても、PDMなどのマークをつけてそれが分かるように明示することがよい。

(5) 活用のためのコミュニティ形成

(活用コミュニティの形成)

分野・地域のコミュニティのつなぎ役は、情報間の関連付けで示したような用語の統制やメタデータの標準化等の取組(4(3)に既述)が求められるが、加えて、活用を進めるための新たなコミュニティの形成にも寄与することが望ましい。データの活用者自らがどのように利用し活用したのかを伝えることができる、事例共有の場を設定できるとよい。

デジタルアーカイブは構築して終わりではなく、活用を促す取組は継続して行われる必要がある。個々のアーカイブ機関の努力も必要であるが、つなぎ役がデータをつなぐことに加えて、人と人のつながりを生み出すことも求められる。

(活用コミュニティの活性化のために)

つなぎ役やアーカイブ機関は、デジタルアーカイブのデータ活用を促進する各種イベントの実施や、バーチャルでの発表の場の設定など、活用コミュニティを活性化するための取組が求められる。例えば、文化財の土器の3Dデータを使ってクリエイターがインテリアやアクセサリーのデザインに取り込むといった活用を想定した場合、全く業種の異なる人たちが集まるイベントでデジタルアーカイブを紹介することも考えられるし、データを使って生み出されるものに表彰するといった仕組みの用意も考えられる。

また、イベント等では、活用者が使いやすいよう、つなぎ役やアーカイブ機関は、提供するメタデータに関する解説や、応用の際のヒントになる情報など、データの活用につながる情報を積極的に発信することが求められる。活用者も、その活用コミュニティに参加し、情報提供を行うことが望まれる。

おわりに

(本ガイドラインの改訂)

デジタルアーカイブを取り巻く環境は変化しつつある。こうした変化に対応するために、本ガイドラインの改訂も必要になる。そのためには、関係省庁、分野・地域のコミュニティのつなぎ役、主要アーカイブ機関、アーカイブの活用者、有識者といったデジタルアーカイブの関係者が一堂に会するフォーラムを設置する必要がある。また、そのフォーラムでは、我が国全体のリソースが効率的に重複なく配分されるよう目配りすること、国の分野横断統合ポータル¹の運営に関して分野を横断するメタデータの標準化の検討、更には、活用領域における活性化のための取組の検討といったことも求められる。

(本ガイドラインの普及)

国は、本ガイドラインの普及に関しては、分野・地域のコミュニティへの直接的な働きかけを行うとともに、ガイドラインに沿った活動を行っている機関・団体・個人を評価し、助成・支援を行う仕組みを構築することが望まれる。こうした取組は、上述のフォーラムが中心となってフォローアップしていくことが求められる。

(我が国におけるデジタルアーカイブの推進のために)

これまで、行政府の様々なところでデジタルアーカイブの構築について検討され、各種のアーカイブ機関で多様なデジタルアーカイブが実際に構築されてきた。しかし、それらを有機的につなげて我が国の社会的発展や国際社会への貢献に広げるための国としての全体像はほとんど見えていなかったといえる。今回の会議体では、関係省庁や主要アーカイブ機関が一堂に会し、有識者や関係者の意見を聴取し、我が国での望ましいデジタルアーカイブ社会の議論を行った。本ガイドラインは、その望ましい社会の実現に向けて、それぞれが何を行うことが求められているのかについて、具体的にまとめたものである。このことは、既述してきたことであるが、最後に改めて述べておく。

本ガイドラインが、我が国におけるデジタルアーカイブの一層の推進と、デジタルアーカイブの更なる活用の促進に資するものとなることを期待している。

(補足資料)

















●用語集

本ガイドラインでは、以下の意味として扱う。

用語	意味
アーカイブ機関	アーカイブ/アーカイブズは、公文書館等を指す言葉として一般的に用いられてきたが、ここでは広い意味での記録機関全般を指すこととする。社会・文化・学術情報資源である資料・作品等のコンテンツを収集し、その資源を整理(組織化)し、保存し、提供する機能を持つ機関・団体等をいう。博物館・美術館、図書館、文書館といった文化的施設のほか、大学・研究機関、企業、市民団体、官公庁・地方公共団体等を含む。なお、提供機能が限定的であり、一般への公開を想定していない機関等も含む。
活用者	デジタルアーカイブ上の様々なデータを活用する者。「アーカイブ機関」が自らのために活用することに加え、一般ユーザ、IT 技術者、クリエイターなど、様々な機関・団体・個人がデータの活用者になりうる。データがオープンな場合、電子展示会の開催、観光用 VR のアプリ提供、教育目的でのコンテンツ利用、防災ポータル構築、人工知能(AI)の学習用データ又は研究対象のデータとしての活用など、様々な目的での活用方法が考えられる。
デジタルアーカイブ	ここでは、様々なデジタル情報資源を収集・保存・提供する仕組みの総体をいう。デジタルアーカイブで扱うデジタル情報資源は、「デジタルコンテンツ」だけでなく、アナログ媒体の資料・作品も含む「コンテンツ」の内容や所在に関する情報を記述した「メタデータ」や、コンテンツの縮小版や部分表示である「サムネイル/プレビュー」も対象とする。
デジタルコンテンツ	アナログ媒体の資料・作品等をデジタル化した高品質なデジタルコンテンツ、あるいはポーンデジタルの作品(デジタルカメラの写真、電子書籍等)であればそのデジタルコンテンツ自体をいう。
コンテンツ	個々の社会・文化・学術情報資源をいう。「デジタルコンテンツ」の他、アナログ媒体の資料・作品等を含む。
メタデータ	データに関するデータを意味し、「サムネイル/プレビュー」や用語を統制するための語彙等も含むものとして広く定義されることもあるが、ここでは、コンテンツの内容、外形、所在等に関する記述等のデータをいう。図書館における書誌データ、アーカイブ機関の収蔵品等の目録データ、文化財の基礎データ等のテキストデータや URI 参照のデータを指す。
サムネイル/プレビュー	コンテンツの要約又は一部分の表示。縮小した画像(サムネイル)、本文テキストの一部表示や数秒程度の音声・動画(プレビュー)等をいう。
国の分野横断統合ポータル	我が国が保有する様々なコンテンツのメタデータに関し、分野横断的な検索を可能とするポータルサイト。「知的財産推進計画 2015」において、その構築の必要性が掲げられ、「知的財産推進計画 2016」では、国立国会図書館サーチがこの分野横断統合ポータルとして位置付けられた。第 4 回実務者協議会では、国立国会図書館は国立国会図書館サーチを発展させた「ジャパンサーチ(仮称)」構想を発表している。

用語	意味
つなぎ役	Europeana の「アグリゲーター」、DPLA の「ハブ」に相当する役割・機能を果たす機関のことをいい、分野・地域コミュニティにおけるメタデータを集約し、API 等による提供を行う機関。メタデータの標準化、用語の統制（辞書・典拠・シソーラスの管理）を行い、コミュニティにおけるメタデータ整備やデジタルコンテンツ作成への支援、長期アクセス保証のための基盤提供、活用コミュニティの形成などの役割を担う。単独の組織が担う場合もあれば、行政、NPO や大学等と分担して行う場合もありうる（詳しくは、「我が国におけるデジタルアーカイブ推進の方向性」を参照のこと。）。
オープン化	インターネットを通じて広く公開されており、商業利用も含めて、目的に応じた活用可能な条件が明示され、手続を要せずに提供されるデータが第三者に自由に利用できるようになっている状態をいう。なお、本ガイドラインは、PDM、CC0、CC BY、特に CC4.0 の利用条件での公開を推奨している。
パブリック・ドメイン	著作権等による保護がなく、自由に利用可能な状態。著作権の保護期間の終了によって、著作権等による制限がなくなった場合や、著作物ではない場合など。
政府標準利用規約	電子行政オープンデータ実務者会議の議論等を踏まえて作成された各府省ウェブサイトの利用規約のひな形。2014 年 6 月 19 日に各府省 CIO 連絡会議で第 1.0 版が決定された。よりデータの活用が進む環境作りに向けて、2015 年 12 月 24 日の各府省 CIO 連絡会議において、第 1.0 版の改定版となる第 2.0 版が決定された。第 2.0 版は CC BY4.0 と互換性をもつ。
API	Application Programming Interface。オペレーティングシステムやアプリケーションソフトが、他のアプリケーションソフトに対し、機能の一部を利用できるように提供するインターフェースのことをいう。
Linked Data(LD)	ウェブ上でのデータの共有・利用を促進するための技術の総称 そのデータが自由な二次利用が可能である場合は、Linked Open Data (LOD)という。
RDF	Resource Description Framework。異なるアプリケーションソフト同士が、組織や領域を超えてデータ交換を可能にするためのデータ記述の基本的な方法のことをいう。
SPARQL endpoint	RDF データを格納したデータベースに対して、標準的な方法（SPARQL クエリ）で検索を可能にするサービスのことをいう。

●利用条件表示一覧

利用条件の表記方法	表記の意味
パブリック・ドメイン・ツール	
CC0	  <p>著作権法上認められる、その者が持つすべての権利(その作品に関する権利や隣接する権利を含む。)を、法令上認められる最大限の範囲で放棄して、パブリック・ドメインに提供すること意味する。</p>
パブリック・ドメイン・マーク	  <p>著作権による制限がなく、自由に利用可能であることを意味する。</p>
クリエイティブ・コモンズ・ライセンス	
※バージョンによって細かい利用条件があることに留意が必要。本ガイドラインでは CC4.0 を推奨。	
CC BY (表示)	  <p>原作者のクレジット(氏名、作品タイトルなど)を表示することを主な条件とし、改変はもちろん、営利目的での二次利用も許可される最も自由度の高いライセンス。</p>
CC BY-SA (表示-継承)	  <p>原作者のクレジット(氏名、作品タイトルなど)を表示し、改変した場合には元の作品と同じライセンス(このライセンス)で公開することを主な条件に、営利目的での二次利用も許可されるライセンス。</p>
CC BY-ND (表示-改変禁止)	  <p>原作者のクレジット(氏名、作品タイトルなど)を表示し、かつ元の作品を改変しないことを主な条件に、営利目的での利用(転載、コピー、共有)が行えるライセンス。</p>
CC BY-NC (表示-非営利)	  <p>原作者のクレジット(氏名、作品タイトルなど)を表示し、かつ非営利目的であることを主な条件に、改変したり再配布したりすることができるライセンス。</p>
CC BY-NC-SA (表示-非営利-継承)	  <p>原作者のクレジット(氏名、作品タイトルなど)を表示し、かつ非営利目的に限り、また改変を行った際には元の作品と同じ組み合わせのライセンスで公開することを主な条件に、改変したり再配布したりすることができるライセンス。</p>
CC BY-NC-ND (表示-非営利-改変禁止)	  <p>原作者のクレジット(氏名、作品タイトルなど)を表示し、かつ非営利目的であり、そして元の作品を改変しないことを主な条件に、作品を自由に再配布できるライセンス。</p>

(出典)クリエイティブ・コモンズのウェブサイト<<https://creativecommons.org/>>及びクリエイティブ・コモンズ・ジャパンのウェブサイト<<https://creativecommons.jp/>>に基づき作成。

(参考資料)

●確認すべき標準・ガイドライン等

(デジタルアーカイブの整備に関するもの)

デジタルアーカイブの構築・連携のためのガイドライン(2012年3月26日)	
作成者	総務省
URL	http://www.soumu.go.jp/main_content/000153595.pdf
概要	図書・出版物、公文書、美術品・博物館、歴史資料等について、連携可能なデジタルアーカイブを構築する手順等が示されている。
震災関連デジタルアーカイブ構築・運用のためのガイドライン(2013年3月)	
作成者	総務省
URL	http://www.soumu.go.jp/main_content/000225069.pdf
概要	震災関連の資料・記録のデジタルデータ化、メタデータ作成、コンテンツ公開・保存・活用、権利関係の処理等について、手順と方法をまとめたガイドライン。
文化資源のデジタル化に関するハンドブック(2011年11月1日)	
作成者	東京大学大学院情報学環・凸版印刷株式会社共同研究プロジェクト
URL	http://www.center.iit.u-tokyo.ac.jp/handbook
概要	様々な文化資源をデジタル化して公開する際の標準的なフローを示すとともに、デジタル化に関する標準的な規格を紹介したハンドブック。
メディア芸術データベースガイドライン(2015年3月31日)	
作成者	文化庁
URL	http://www.bunka.go.jp/seisaku/geijutsubunka/media_art/pdf/media_arts_db_guideline.pdf
概要	メディア芸術データベースの構築にあたって得られたノウハウなどの成果が整理され、メディア芸術関係のデータベース・アーカイブなどの事例が示されている。
全国の公文書館等におけるデジタルアーカイブ・システムの標準仕様書(2009年3月作成、2012年10月一部改訂)	
作成者	国立公文書館
URL	http://www.archives.go.jp/law/pdf/da_121210.pdf
概要	公文書館におけるデジタルアーカイブ・システムを構築する方法(目録情報のデータ化・データベース化、資料のデジタル画像化等)がまとめられている。
政府情報システムの整備及び管理に関する標準ガイドライン(2014年12月3日)	
作成者	各府省情報化統括責任者(CIO)連絡会議
URL	http://www.soumu.go.jp/main_content/000352682.pdf
概要	各府省が情報システムを整備し運用する際に、標準的に準拠すべきルール・方針。
政府情報システムの整備及び管理に関する標準ガイドライン実務手引書(2015年3月19日)	
作成者	内閣官房情報通信技術(IT)総合戦略室、総務省行政管理局
URL	http://www.soumu.go.jp/main_content/000348369.pdf
概要	「政府情報システムの整備及び管理に関する標準ガイドライン」の内容の解説や参考となる事例等が掲載されている。
資料デジタル化の手引(2017年版)	
作成者	国立国会図書館
URL	http://www.ndl.go.jp/jp/aboutus/digitization/guide.html
概要	図書館資料のデジタル化の方法・仕様、画像データの品質検査等について解説。

Reference Model for an Open Archival Information System(OAIS 参照モデル)	
作成者	宇宙データシステム諮問委員会(CCSDS)
URL	https://public.ccsds.org/pubs/650x0m2.pdf https://www.iso.org/standard/57284.html
概要	デジタル情報の長期保存アーカイブシステムに関する枠組みを規定した国際標準規格(ISO 14721:2012)。 (参考)栗山正光「デジタル情報保存のためのメタデータに関する動向」 http://current.ndl.go.jp/ca1489 (参考)Brian Lavoie「The Open Archival Information System (OAIS) Reference Model : Introductory Guide (2nd Edition)」 https://doi.org/10.7207/twr14-02

(メタデータの標準に関するもの)

メタデータ情報共有のためのガイドライン(2011年3月31日)	
作成者	筑波大学ほか
URL	http://www.mi3.or.jp/item/A03.pdf
概要	メタデータの提供者、活用者双方を対象に、メタデータ的设计、作成から利用、運用管理まで、メタデータの相互運用性、長期利用可能性を高めるための指針。
ミュージアム資料情報構造化モデル	
作成者	東京国立博物館
URL	http://webarchives.tnm.jp/docs/informatics/smmoi/
概要	博物館・美術館における歴史・民俗・考古・美術の各種資料がもつ情報とその記述のモデル。
国立美術館所蔵作品総合目録検索システム作品詳細・凡例	
作成者	国立美術館
URL	http://search.artmuseums.go.jp/rai.html
概要	独立行政法人国立美術館の4つの美術館の所蔵作品の総合目録検索システムにおける「作品詳細」画面の凡例。
国立国会図書館ダブリンコアメタデータ記述(DC-NDL)	
作成者	国立国会図書館
URL	http://www.ndl.go.jp/jp/aboutus/standards/meta/about_dcndl.html http://www.ndl.go.jp/jp/aboutus/standards/meta/dcndl_examples.html
概要	国立国会図書館がインターネット上に存在する情報資源等の組織化・利用提供のために定めたメタデータ記述語彙及び記述規則。
FOAF Vocabulary Specification	
作成者	Dan Brickley, Libby Miller
URL	http://xmlns.com/foaf/spec/
概要	RDFを使用して人とその活動に関する情報を記述するために用いられる語彙。
SKOS: Simple Knowledge Organization System	
作成者	World Wide Web Consortium (W3C)
URL	https://www.w3.org/2004/02/skos/ http://www.w3.org/TR/skos-reference
概要	シソーラス、分類体系、件名標目表などの統制語彙を表現するための標準的なデータモデルと語彙。

CDWA: Categories for the Description of Works of Art	
作成者	ゲティ財団
URL	http://www.getty.edu/research/publications/electronic_publications/cdwa/
概要	美術作品、建築物等を記述するためのガイドライン。CDWA に基づくメタデータ交換のための XML スキーマとして CDWA Lite がある。 http://www.getty.edu/research/publications/electronic_publications/cdwa/cdwalite.pdf
CIDOC CRM: CIDOC Conceptual Reference Model(CIDOC 概念参照モデル)	
作成者	国際博物館会議(ICOM) 国際ドキュメンテーション委員会(CIDOC)
URL	http://www.cidoc-crm.org/
概要	博物館の収蔵品に関するドキュメンテーションの標準。ISO 21127:2014 として国際標準化されている。 (参考)村田良二「CIDOC CRM モデリング入門」 http://ryoji.sakura.ne.jp/museuminfo/crm-model-intro/
Darwin Core	
作成者	Biodiversity Information Standards (TDWG)
URL	http://rs.tdwg.org/dwc/
概要	生物多様性情報を共有するために開発されたメタデータ標準。生物の標本、観察データを交換するために使用される。
ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families	
作成者	国際公文書館会議(ICA)
URL	http://www.ica.org/en/isaar-cpf-international-standard-archival-authority-record-corporate-bodies-persons-and-families-2nd
概要	文書館資料の作成者の典拠レコード情報を記述するための標準。
ISAD(G): General International Standard Archival Description	
作成者	国際公文書館会議(ICA)
URL	http://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition
概要	文書館の資料を記述するための標準。
EAD: Encoded Archival Description	
作成者	米国アーキビスト協会(SAA)
URL	https://www.loc.gov/ead/
概要	文書館の検索手段をオンラインで利用可能にするための標準。
ISIL: International Standard Identifier for Libraries and Related Organizations(図書館及び関連組織のための国際標準識別子)	
作成者	国際標準化機構(ISO)
URL	http://biblstandard.dk/isil/index.htm http://www.ndl.go.jp/jp/library/isil/
概要	図書館をはじめ博物館、美術館、文書館などの類縁機関に付与される国際的な識別子。ISO 15511:2011 として国際標準化されている。図書館や博物館、美術館などのサービスや、コンテンツ・流通ビジネスなどで自由に利用可能。国立国会図書館が日本における国内登録機関であり、日本国内の ISIL 付与及び管理を行う。

DOI : Digital Object Identifier (デジタルオブジェクト識別子)	
作成者	国際 DOI 財団
URL	http://www.doi.org/ http://www.doi.org/doi handbook/translations/japanese/hb.html
概要	コンテンツの電子データに付与される国際的な識別子。ISO 26324:2012 として国際標準化されている。デジタルネットワーク上で管理される情報の相互運用可能な交換と永続的で作動可能な識別のためのシステムを提供する。DOIの登録には国際 DOI 財団の認める DOI 登録機関への参加が必要である。ジャパンリンクセンター (JaLC) は、日本における唯一の DOI 登録機関であり、科学技術振興機構、物質・材料研究機構、国立情報学研究所及び国立国会図書館によって、運営されている。 https://japanlinkcenter.org

(オープン化に関するもの)

政府標準利用規約(第 2.0 版)(2015 年 12 月 24 日決定)	
作成者	各府省情報化統括責任者(CIO)連絡会議
URL	http://www.kantei.go.jp/jp/singi/it2/densi/
概要	各府省ウェブサイトの利用規約のひな形。CC BY と互換性があることを明記するなどのルールが示されている。
二次利用の促進のための府省のデータ公開に関する基本的考え方(ガイドライン)(2013 年 6 月 25 日決定 2015 年 12 月 24 日改定)	
作成者	各府省情報化統括責任者(CIO)連絡会議
URL	http://www.kantei.go.jp/jp/singi/it2/densi/kettei/data/gl27_honbun.pdf
概要	各府省の保有するデータの公開に関する基本的考え方を整理したガイドライン。公開できない理由が明確なものを除き、保有するデータは全て公開するという理念が示されている。
(二次利用の促進のための府省のデータ公開に関する基本的考え方(ガイドライン)の別添 2) 数値(表)、文章、地理空間情報のデータ作成に当たっての留意事項	
作成者	各府省情報化統括責任者(CIO)連絡会議
URL	http://www.kantei.go.jp/jp/singi/it2/densi/kettei/data/gl26_betten2.pdf
概要	上記ガイドラインの別添資料。数値(表)、文章、地理空間情報について、機械判読が容易なデータを作成するに当たっての留意事項が示されている。
地方公共団体オープンデータ推進ガイドライン(2015 年 2 月 12 日)	
作成者	内閣官房情報通信技術(IT)総合戦略室
URL	http://www.kantei.go.jp/jp/singi/it2/densi/kettei/opendata_guideline_r4.pdf
概要	地方公共団体におけるオープンデータの推進に係る基本的考え方等を整理し、地方公共団体がオープンデータに取り組むに当たっての参考として作成された。
我が国におけるオープンサイエンス推進のあり方について～サイエンスの新たな飛躍の時代の幕開け～(2015 年 3 月 30 日)	
作成者	国際的動向を踏まえたオープンサイエンスに関する検討会
URL	http://www8.cao.go.jp/cstp/sonota/openscience/index.html
概要	公的研究資金による研究成果のうち、論文及び論文のエビデンスとしての研究データは原則公開とする方針が示されている。

独立行政法人国立公文書館における公文書管理法に基づく利用請求に対する処分に係る審査基準(2011年4月1日決定 2013年4月1日改訂)	
作成者	国立公文書館
URL	http://www.archives.go.jp/information/pdf/riyoushinsa_2011_00.pdf
概要	別添参考「30年を経過した特定歴史公文書等に記録されている個人情報について」があり、公文書館における情報公開において、情報の機微性に応じ一定期間を経て公開する措置を取る際の基準が示されている。

(データの共有方法に関するもの)

みんなの公共サイト運用ガイドライン(2016年版)	
作成者	総務省
URL	http://www.soumu.go.jp/main_content/000439213.pdf
概要	高齢者や障害者を含めて、誰もが支障なく利用できるよう、アクセシビリティを確保するためのガイドライン。

(国立国会図書館サーチの)外部提供インターフェース(API)	
作成者	国立国会図書館
URL	http://iss.ndl.go.jp/information/api/
概要	国立国会図書館サーチが提供する、メタデータ連携のための API についての説明。英語版の仕様書もある。同「国立国会図書館サーチが提供する OAI-PMH」のページには、OAI-PMH の概要と実装時の注意点などをまとめた「OAI-PMH の要点」(ZIP ファイル)が公開されている。 http://iss.ndl.go.jp/information/api/oai-pmh/info/

「ロメジュリ」って何? : 識別子を巡って『情報管理』Vol.59, No.11(2017年2月)	
作成者	ゼノン・リミテッド・パートナーズ 神崎正英
URL	http://doi.org/10.1241/johokanri.59.772
概要	Linked Data において重要な役割を果たす URI について、わかりやすく解説されている。

Linked Data	
作成者	The World Wide Web Consortium (W3C)
URL	https://www.w3.org/standards/semanticweb/data
概要	Linked Data の概要、用途、使用例等が紹介されている。

RDF 1.1 Primer	
作成者	Guus Schreiber and Yves Raimond (eds.)
URL	https://www.w3.org/TR/rdf11-primer/
概要	RDF の基本的な概念や具体的な使用例など、RDF を効果的に使用するために必要な知識について、例を用いてわかりやすくまとめた入門用文書。

Web NDL Authorities SPARQL API 仕様書(2014年3月31日)	
作成者	国立国会図書館
URL	https://id.ndl.go.jp/information/wp-content/uploads/2014/03/api-spec.pdf
概要	SPARQL クエリの仕様を解説し、それを用いて Web NDL Authorities の典拠情報(RDF 形式のデータ)を検索する方法等が示されている。

IIIF : International Image Interoperability Framework	
作成者	IIIF Consortium
URL	http://iiif.io/
概要	<p>デジタル画像の共有において相互運用可能性を担保するための仕様。標準的なウェブ技術を用いて、デジタル画像の解析と共有、システム間の移行、高度な画像アクセスを容易にする。複数のデジタルアーカイブが提供する画像を、同一のビューワーで閲覧することが可能になる。</p> <p>(参考)永崎研宣「デジタル文化資料の国際化に向けて:IIIFとTEI」『情報の科学と技術』67(2), p.61-66, 情報科学技術協会. http://doi.org/10.18919/jkg.67.2.61</p>

●活用できる表形式のデータとは？

エクセル等のアプリケーションソフトを用いた表形式のデータを作成する際は、機械可読性を担保するため、以下の点に留意して作業を行わなければならない。

- 表データは、列を項目、行をレコードとして構成する。
- 1つのデータセルには、1つの要素のみ記入する。
- データセルに、整形や位取りのための文字(スペース、改行、数値でのカンマ等)を含めない。
- 数値等のデータの値やタイトル、単位以外の情報を、セルに含めない。
- 値が存在しない場合を除き、データセルを空白にしない。(データ値を省略しない)
- 英数字は半角とし、ひらがな・カタカナは全角とする。
- データ列の内容を示す項目名は、1行で構成する。
- レイアウトのための空行・列は使わない。
- データセルにふりかな、コメント・注釈などの加工は行わない。
- 全てのセルは、他のセルと結合しない。
- ファイルのヘッダやコメント等の内容が残っていないか、プロパティ情報(ファイルの属性情報)に問題がないか十分に確認してから公開する。

(参考)「数値(表)、文章、地理空間情報のデータ作成に当たっての留意事項」(二次利用の促進のための府省のデータ公開に関する基本的考え方(ガイドライン)別添 2(2013年6月25日各府省情報化統括責任者(CIO)連絡会議決定 2015年12月24日改定))
http://www.kantei.go.jp/jp/singi/it2/densi/kettei/data/gl26_betten2.pdf

●よくある質問

〔メタデータの整備とデジタル化に関すること〕

Q:そもそも、アーカイブ機関にとってデジタルデータを作成するメリットは何ですか？

A:メタデータを電子的に整備したり、コンテンツをデジタル化するなどしてデジタルデータを作成する最大のメリットは、「時間や場所を問わず利用が可能になること」です。自館の業務効率化はもちろん、サービスの利便性を向上させ、より魅力あるものにすることができます。さらに、ウェブ上に公開することによって「世界中の多くの人々に知ってもらえること」です。多くの人に知られることで、本物の資料・作品への理解を深めてもらう契機となり、実物へのアクセスが増える、アーカイブ機関の価値が高まるなどの効果が期待されます。

Q:なぜメタデータを整備しなければいけないのでしょうか？

A:上記のようなメリットを十分に活かすためには、「検索で見つけてもらえること」が最大のカギとなります。その際、アーカイブ機関がどのようなコンテンツを持っているかという情報を適切に提供することが重要です。簡単なメタデータだけでは、目的や内容から検索することが難しくなります。様々な「つなぎ役」と連携し、より広く活用してもらうためには、より整備されたメタデータが求められます。また、わかりやすい管理的なメタデータは、データを長期に保存し運用するためにも必要です。

〔オープン化に関すること〕

Q:デジタルデータを公開してオープンにすることで来館者が減ったりしませんか？

A:デジタル化してウェブ上に公開することで、むしろ来館者が増えることが多いと言われています。ウェブ上で見て満足するよりは、そこで得た知見を契機に資料・作品への興味を深めて現地に足を運ぶという、有料来館であっても、ウェブによる宣伝効果の方が大きいようです。また、公開の仕方によっては、画像の提供や作品サイズ等の確認といったことが可能になるため、データ提供にかかっていた人的コストの削減につながるというメリットもあります。

Q:デジタル化し、公開したら、問い合わせが増えて大変ではありませんか？

A:デジタル化することによって活用が広がることは、所蔵する資料・作品の価値を高めることにつながります。また、問い合わせの実績は、アーカイブ機関の資料・作品の価値評価にもつながりえますので、問い合わせ件数がメリットになるとも考えられます。

Q:データをオープンにして、改変されてしまったら、資料・作品の価値が下がるのでは？

A:基本的に、自分たちの機関が真正性の高いデータを保持し、他機関との共有や公開を行っていれば、不自然に改変されたデータは淘汰されますし、改変した人物に批判が集まります。言い換えれば、真正性の高いデータを持続的に保持し続ける使命が自らの機関にはあるという根拠にもなります。なお、正当な目的の改変であれば、通常は資料・作品を多く知ってもらう機会になることが多いです。

Q:現在写真データを有償頒布していますが、その売上げが下がりませんか？

A:写真データ頒布にかかる事務などのコストと有償頒布のコストが見合っているかを確認しましょう。オープンにすることで事務処理がなくなり、業務軽減になっている事例が多いと聞きます。また、売上げによる自己収入を評価項目として掲げているところもありますが、自らの価値を高めて良い評価を得るには、オープンデータ・オープンイノベーション・オープンアクセスに貢献することでより広く活用される方がより効果的であると考えます。もちろん、高品質の写真を販売することが、別の原資になる場合には、このガイドラインがそれを妨げるものではありません。

〔システム運用に関すること〕

Q:デジタルアーカイブを運用するためのサーバ機器などを用意するコストが捻出できません

A:データをインターネット上に載せるだけでも、意味は十分にあります。表形式のテキストデータでも、ウェブ上にデータがあれば、「つなぎ役」や「活用者」が検索できるようにしてくれる可能性があります。また、データの発見から共同事業化に発展すれば、サーバを運用している機関との連携も可能になるかもしれません。重要なことは、自分たちの機関がどのような形でライセンスを明確にして、データをオープンにすることです。

Q:ウェブ上にある既存のサービス等で運用するのではだめですか？

A:既存のサービス等を用いてデータをオープンに公開することで、より多くの人々が見られることは意味がありますので、十分に検討に値します。一方で、データの真正性や持続的提供については他者に依存することになります(例えば、データをアップする際には無料だったが、ある時期から提供者が使用料を払わないとデータが見られなくなる等)。独自のサービス展開が行いにくくなるというデメリットがあることも踏まえ、デジタルデータを広く共有させオープンに展開できるよう、本ガイドラインでは、つなぎ役との連携ができる方式も推奨していますので、その点からも検討をされることが望まれます。

Q:自機関でシステムを導入する際の注意点は？

A:第三者が活用しやすい仕様として、データをオープンにする範囲をどの部分までにするかを検討し、ダウンロード等が容易かなどを確認してください。また、データの登録や削除、統計処理等の管理が自分でできるシステムであること、ガイドラインにあるような「つなぎ役」との連携が可能な機能があることが望まれます。

併せて、いわゆる「ベンダーロック(ある特定業者のシステムに依存する状態。業者を変えられずコストが上がり、業者がシステム提供を終えたことでデジタルアーカイブが維持できなくなるなどの弊害がある)」にならないよう留意する必要があります。

●ガイドライン早見表

本ガイドラインにおいて示した、各アーカイブ機関やつなぎ役の機関が取り組むべき事項について、レベルに応じた内容を以下に示す。

	各アーカイブ機関		つなぎ役 ※原則は左列の「望ましいレベル」、加えて以下の機能が望まれる
	必要レベル	望ましいレベル	
メタデータの整備	<ul style="list-style-type: none"> ○独自の統一フォーマットだがガイドラインの必須項目を含む ○表形式のデータ(CSV等)での整備 ○識別子の付与(重複しない管理番号の付与) 	<ul style="list-style-type: none"> ○同じ分野の標準フォーマットを使用 ○自館又はコミュニティ等が提供するシステムによる整備 ○タイトル等の多言語化(英語又はローマ字表記等)対応 (○DOIの付与)※なおよいレベル 	<ul style="list-style-type: none"> ○標準化の推進 ○用語(辞書・典拠・シソーラス)の統制 ○コミュニティ内の機関識別子の付与・管理 ○分野の共通用語へのURI付与。又はWikidata、DBpedia等に識別リンクを提供
サムネイル/プレビューの作成		<ul style="list-style-type: none"> ○サムネイル/プレビューの作成 	
デジタルコンテンツの作成	<ul style="list-style-type: none"> ○カラーチャートやメジャーを参照できる(画像の場合) ○一般的なフォーマットで用意 	<ul style="list-style-type: none"> ○できる限り高品質なものを作成 	
長期アクセスの保証	<ul style="list-style-type: none"> ○メタデータ、コンテンツを定期的に確認し、リンク切れがないように確認 	<ul style="list-style-type: none"> ○自らで安定的なシステム確保、又はコミュニティ等で提供される安定的なシステムを利用 ○自らの場合、システムリプレースの人員・予算の確保 ○自らの場合、データ移行を見込んだ人員・予算の確保 	<ul style="list-style-type: none"> ○安定的なプラットフォームの提供
メタデータの共有方法	<ul style="list-style-type: none"> ○表形式のデータ(CSV等)の用意 	<ul style="list-style-type: none"> ○ハーベスト用API(OAI-PMH)の用意 ○検索用APIの用意(SPARQL含む) 	
メタデータの提供	<ul style="list-style-type: none"> ○表形式のデータのウェブ公開(若しくはつなぎ役への提供)、又は所蔵資料・収蔵品のデータにURIを付与 	<ul style="list-style-type: none"> ○データベースでの公開 ○URIのあるメタデータセットを公開 	<ul style="list-style-type: none"> ○(分野・地域コミュニティの)ポータルでの公開 ○ハーベスト用API(OAI-PMH)の提供 ○検索用APIの提供(SPARQL含む)

	各アーカイブ機関		つなぎ役
	必要レベル	望ましいレベル	※原則は左列の「望ましいレベル」、加えて以下の機能が望まれる
サムネイル/プレビューの提供	○サムネイル/プレビューをウェブ公開	○サムネイル/プレビューを検索結果に表示 ○メタデータにサムネイル/プレビューの URL の埋め込み	
デジタルコンテンツの提供	○デジタルコンテンツにメタデータを関連付けてウェブ公開	○サムネイル/プレビューからデジタルコンテンツへのリンクを用意 ○相互運用可能な方式（画像の場合は IIIF 等）での提供	
オープン化等対応	○メタデータ、サムネイル/プレビュー、デジタルコンテンツへの CC ライセンス付与、又は二次利用条件の表示	○共有（連携）するメタデータの CC0 採用 ○可能なものについてサムネイル/プレビュー及びデジタルコンテンツの CC0 又は CC BY 採用 ○著作権保護期間が満了したサムネイル/プレビュー及びデジタルコンテンツの PDM の表示 ○提供しているデジタルコンテンツの二次利用条件の表示の徹底	
アーカイブデータの活用		○活用コミュニティへのメタデータの解説等の情報提供	○付加価値情報の付与 ○情報間の関連付け（メタデータ整備の項も参照） ○活用コミュニティの形成